人工智能赋能数字政府建设的安全风险及应对策略研究

康京涛 王 砼
（西安财经大学法学院 西安 710061）

摘 要：［目的/意义］开展人工智能赋能数字政府建设的风险分析，研究风险应对策略，推进我国数字政府建设高质量发展。［方法/过程］综合运用文献研究、案例研究等方法，分析了人工智能赋能数字政府建设中在国家、政府、社会、个人四个层面存在的风险隐患。［结果/结论］从推动人工智能技术的分类分级、树立敏捷治理的治理理念、完善多元化的人工智能审查评估机制和明确人工智能在数字政府建设中的责任链条四个方面提出风险应对策略，为我国人工智能赋能数字政府建设的风险治理提供参考。

关键词：人工智能 数字政府建设 数据主权 社会歧视 公民隐私
分类号：D922.1 D630
DOI：10.31193/SSAP.J.ISSN.2096-6695.2024.03.02

0 引 言

推动政府数字化转型，加快建设“数字中国”，是国家治理体系和治理能力现代化的重要支撑，也是回应社会变迁和技术发展双重需求的新型治理理念[1]。ChatGPT等生成式人工智能正在以前所未有的深度、广度、速度进入社会生活，政府业务方式也开始由电子化、网络化、协同化向智能化变革，政府业务形态也从办公自动化、电子政务向智能化的数字政府转变。智能化的数字政府嵌入了人工智能功能的各种政务应用系统，实现了政务应用系统的“智能协同”“智能通办”“智能通管”。智能化的政府运行新形态意味着大模型将成为政务云的基本构成要素，政务数字化应用系统发生智能重构，政务云建设方式将从“云平台（Iaas+Paas）+政务应用系统（Saas）”向“云平台（Iaas+Paas）+大模型（Maas）+政务智能应用系统（ASaas）”转变[2]。以智能化为核心的数字政府建设关键在于运用算法进行自动化行政。自动化行政将行政过程中的事实与规则处理转化为代码自动运行，具有高效、科学、精准的特性。

［作者简介］康京涛，男，教授，博士，研究方向为数字法治政府、数据法治，Email: Kangjingtao2011@126.com；王砼，男，硕士生，研究方向为数字法治政府、数据法治，Email: wangtong2913@126.com（通讯作者）。

目前，人工智能赋能数字政府建设成为全球政府管理创新领域的核心议题。对此，美国成立了专门管理人工智能的机构，欧盟及其成员国也都出台了人工智能发展战略文件。在我国，早在2009年，南京市环保局（已更名为南京市生态环境局）就研发了“环境保护行政处罚自由裁量辅助决策系统”；2014年，天津市人大常委会法制工作委员会使用北大法宝智能立法支持平台辅助地方立法；2017年，广西壮族自治区上线“广西智能审批和审管信息一体化系统”，实现了12种行政审批事项的智能审批与监管；2020年，上海市在“一网通办”虹口频道开通线上无人干预自动办理模块，利用大数据、人工智能等技术，实现了审批自动化、服务便利化。不止于此，人工智能在我国还被应用在交警非现场执法、公共信用评价、行政给付等场景中。毋庸置疑，人工智能嵌入数字政府建设有效提高了政府治理能力，优化了数字政府治理结构。与此同时，人工智能这项技术本身还未发展成熟，具有极大的不确定性。作为一项为数字政府建设赋能的技术，如若人工智能未带来“正效应”，反而暗含算法黑箱、数据主权受侵害、行政脱离法治架构等一系列风险隐患，将会极大程度上削弱公众信任甚至行政机关内部的积极性。因此，有必要对人工智能赋能数字政府建设的风险及其应对策略进行探讨，以期实现人工智能与数字政府的有机结合与良性互动。

1 研究综述

关于人工智能赋能数字政府建设的风险问题，现有研究主要从传统人工智能和生成式人工智能两个视角，围绕风险隐患及其风险应对两个方面开展研究与探讨。

针对人工智能赋能数字政府建设的安全风险，从传统人工智能的视角来看，Calo[3]于2017年通过对人工智能政策的研究，认为人工智能技术的应用将会引发公正和公平、武力使用、安全和认证、隐私和权力、税收和劳动力的转移等问题；Jane Fountain立足数字政府建设中的算法系统运用，认为计算算法面临着算法歧视带来的系统性偏误问题[4]，包括了面部识别技术、预测性警务、公共管理中的自动化决策等问题；汪太贤等[5]从安全、责任、公正、自主四个维度出发，认为人工智能赋能数字政府建设的风险隐患包括数据共享的不安全性、治理责任的模糊化、算法公正的减损化、技术依赖的自主性损害四个方面；李良成等[6]认为人工智能若背离正确价值导向，将在治理场景中引发技术风险、治理风险、伦理风险；刘玮等[7]认为人工智能等技术在政务服务场域中存在着责任体系失衡、主体权力冲突、公共价值迷失等问题。从生成式人工智能的视角来看，隐患类型可分为两类：一是以算法为基点的算法透明可解释风险、算法公平可靠风险、算法安全可控风险、算法问责风险[8]；二是数据主权风险、资本侵蚀风险、信息失序风险[9]。整体而言，已有关于人工智能赋能数字政府建设安全风险研究具有客观性，但受研究者视角不同，给出的安全风险因素存在差异，尚缺乏针对性和系统性。人工智能赋能数字政府建设的安全风险是一个系统性问题，因此本文从国家、政府、社会、个人视角综合考虑，有助于更科学、精准地判定风险隐患。

关于人工智能赋能数字政府建设的风险应对，从传统人工智能的视角来看，Autor[10]、Turner[11]、Wright[12]等从法律和政策角度提出政府需要进一步完善法律和政策以规范人工

智能的发展，重点构建人工智能监测影响评估体系、重点保护公民的信息与数据安全；Urs等[13]立足人工智能引发的信息不对称、寻找规范共识、政府责任不匹配三个挑战，从技术、伦理、法律和社会层面提出了人工智能治理模型；秦小建等[14]提出以公平与安全理念规训算法逻辑、以算法解释义务增进权力透明度以及保持人类智能（Human Intelligence）与人工智能的协同跟进的治理路径。就生成式人工智能的视角来看，曾宇航等[15]认为应从研发数字安全防御技术、全面升级社会安全响应制度、调适优化公共安全处置行动三个方面应对安全风险；张欣等[16]认为应从完善行政决策的监督机制、技术保障机制、主体机制三个方面应对ChatGPT辅助行政决策的算法危机；刘绍宇[17]从技术、制度、权利、应用四个方面提出了研发和推广拥有自主知识产权的ChatGPT模型，将数据分级作为ChatGPT模型的基准和规制路径，搭建ChatGPT模型的责任分配链条和溯因机制，以及构筑ChatGPT模型应用服务的原则体系等构想。总体上，现有研究主要从法律、制度、管理、技术四个层面探讨风险隐患治理路径。

综合人工智能赋能数字政府建设风险隐患及其风险应对的相关研究发现，当前我国人工智能赋能数字政府建设中面临的风险问题已成为学界关注的热点，研究也取得了显著进展。但现有研究仍然存在以下拓展空间：一是现有研究大多立足于单一的政治学、行政学、管理学、法学等视角，对于风险隐患的分析存在一定的学科局限，需要从系统论的视角整体考察；二是现有研究只是笼统地分析风险隐患的来源、类型及其防控措施，较少考虑风险隐患的侵害主体。厘清人工智能赋能数字政府建设风险隐患的侵害主体或者以主体视角分析风险隐患是对其治理的基础。鉴于此，本文遵循人工智能赋能数字政府建设涉及的主体，从国家、政府、社会、个人四个维度出发，分析人工智能赋能数字政府建设的安全风险，进一步探讨风险应对的策略，以期推进数字政府建设高质量发展。

2　人工智能赋能数字政府建设的风险分析

人工智能嵌入数字政府建设有助于修补决策者理性思维上的缺陷，进而建立信息灵通、明达参与、充分讨论的决策过程，是当代公共服务系统理性化和现代化的重要推力[18]。但科学技术固有的“双刃剑”性质使得人工智能赋能数字政府建设中，在国家、政府、社会、个人层面存在着潜在的风险。

2.1　国家层面：数据主权受到威胁

由于人工智能技术发展尚不成熟，人工智能赋能数字政府建设过程中存在数据主权受到威胁的风险。

其一，国家秘密泄露的风险。人工智能作为以数据为核心要素的技术，在应用于数字政府建设中时，可能会涉及我国政治、经济、社会、生态等领域的关键信息和重要数据，一旦数据处理过程出现纰漏，极有可能造成国家秘密泄露，对国家安全构成不可估量的威胁。

其二，遭遇技术霸权的风险。AI基础技术是人工智能的根基，我国在人工智能技术研发中仍有较大的发展空间，在数字政府建设中还存在一定的技术空白点。一方面，容易被数据霸权国

家形成技术垄断与数据监控；另一方面，也易因技术瓶颈限制数字政府建设进程。目前 Open AI 尚未对我国开放，部分国外企业在技术领域占据垄断地位，因此在我国尚未厘清人工智能技术底层逻辑的情况下“贸然”将其应用于数字政府建设，可能遭遇一定的技术困境。

其三，国际合作机会减少的风险。人工智能技术被各国广泛应用于数字政府建设，但由于价值理念、基本诉求等差异，尚未形成统一的数据安全治理机制，达成的数据安全治理机制多为单边、双边、多边框架和贸易规则[19]。在此背景下，我国与其他国家合作空间可能会被压缩，进而导致我国促成全球数字空间治理机制形成并在其中发挥主导作用面临着严峻的国际挑战[20]，数据主权也可能受到相应威胁。

2.2 政府层面：行政权公共性面临冲击

公共性是行政权的合法性基础，因此，行政权的行使必须以维护和促进公共利益为必要目的，否则就会脱离公共行政的主色调。人工智能嵌入数字政府建设面临一定的公共性伦理挑战和对政府行政权公共性冲击的风险。

首先，合理行政遭遇阻却。合理行政是行政裁量的重要理论，其理论结构包含行政行为的目的具有正当性、行政行为的作出必须考虑相关因素、行政行为造成的结果不得明显不当。人工智能技术的本质是依靠海量的数据与强大的算法得出结论，看似是可靠的经验总结，但与人的决策存在天壤之别。传统的以公务人员为主体的行政决策往往会考虑社会规范、社会影响等因素。如近年来我国政府一直倡导的“人性化执法”，就是要求行政机关在作出决策和行政行为时需要综合考量各方面因素，以保证法律效果与社会效果的统一。以人工智能为核心的自动化行政本质上属于“格式化”行政，它如同冰冷的机器一样无法顾及现实生活中的各种因素，进而有可能出现机械行政、机械执法。

其次，程序正当受到挑战。程序正当承载着程序正义的价值。在人工智能赋能数字政府建设的过程中，一方面，由于“算法黑箱”和算法可解释性的缺乏使得公民无法与智能系统有效沟通，公众难以参与决策过程；另一方面，人工智能系统一般由行政机关或其委托的第三方设计，公众只能看到预测结果，无法看到过程，即使行政机关公布了算法的过程，公众也往往因其专业性、复杂性无法理解智能化软件的逻辑、功能、后果，使行政公开流于形式，进而导致公众的陈述、申辩等程序性权利被剥夺。如交通领域的非现场执法受制于违章信息告知的滞后性，公众的陈述权、申辩权无法得到保障。

最后，追责链条发生断裂。传统行政模式中，行政治理的主体是人，形成的是“公务人员—行政相对人”单向度的直接关系，行政机关对自己做出的行为负责并承担相应的责任，追责路径简单明了。依靠人工智能形成的自动化决策使得这种关系转化为“公务人员—自动化设备—行政相对人”。这在一定程度上削弱了人在公共行政中的地位，间接地使人工智能（自动化设备）成为数字政府治理的主体，弱化了数字政府治理中行政行为的合法性。如若自动化决策的结果出现纰漏，损害了行政相对人的信赖利益，造成了经济损失，这一责任由谁承担，损失由谁赔偿，挑战了传统行政中“谁作出谁负责”的追责机制，成为了现代行政法治的难题[21]。

2.3 社会层面：技术不成熟加剧不公平与歧视

人工智能是基于数据和算法的技术系统，随着人工智能和数字政府建设的深度融合，数据和算法自身的缺陷与不足将带来不公平与歧视加剧的风险。

一是数据偏差。人工智能系统正常运转依赖于数据质量，即数据是否均衡、是否具有代表性、是否充足，否则将导致人工智能系统运转出现偏见。换言之，数据一旦出现偏差，算法的结果极易造成歧视。如谷歌公司的图片识别软件曾因数据样本不足、不全面，将黑色人种识别为“大猩猩”。尽管人工智能系统在数据使用之前会进行数据处理和清洗，但局限于有些数据时间跨度长，仍然无法保障数据质量的百分之百准确和无误。人工智能作为一项数据驱动的技术，数据的隐藏、未知和不精确，可能会导致结果偏差和潜在偏见。

二是算法歧视。人工智能决策的过程是数学方法和算法代码集体运行的行为判断集合过程。如若出现数据偏差和技术设计偏误，歧视难以避免。关于数据偏差前文已述，对于技术设计偏误，这是因为算法是由设计者编辑而成的，体现的是设计者的主观意愿。由于设计者的人为目的或主观判断不充分等因素，在客观上无法保证其精准性。如 2018 年美国纽约州被迫终止了防止家庭暴力儿童保护预测算法系统的使用，原因在于算法程序根据数据分析认为一些父母具有“严重的家暴倾向”，致使上万正常父母和子女被迫分离[22]。目前在数字政府建设领域由人工智能系统造成的歧视主要包括两种类型：第一种是偏见代理的算法歧视，如近期备受关注的人工智能 ChatGPT 有可能存在较强的美国地缘色彩，在回答问题、生成答案时可能会偏向于美国利益；第二种是特征选择的算法歧视，如美国司法系统实行的审前释放评估系统（COMPAS），在选择变量生成的风险评估分数中，黑色人种的分数往往要高于白色人种的分数[23]。

2.4 个人层面：隐私遭遇泄露与侵犯

其一，技术层面导致的隐私泄露风险。人工智能赋能数字政府建设的核心是技术赋能，而政府本身不可能生成技术，因而政府多采用公私合作、购买服务等方式获取人工智能技术。在技术的加持下，政府将公众的个人住所、电话号码、家庭情况、健康状况、房产以及生物信息集聚起来，然后进行数据整合和分析，将相关信息运用到决策。这一过程中，由于处理的数据蕴含极大的经济价值与社会价值，在数据处理、结果储存等环节中可能存在一定公众数据泄露的风险。如新冠肺炎疫情期间，一些患者因信息泄露遭到了他人短信和电话的无理谩骂和骚扰，给患者本人和家庭带来了极大不便[24]。以算法为核心的人工智能技术在数据分析和决策中打破了物理障碍，具有“望远镜”和“雷达”功能，个人隐私可能存在泄露的风险。质言之，“技术巨变重构了社会图景，在空间、时间与社会维度引发隐私保护的深刻困境”[25]。

其二，制度层面导致的隐私侵犯风险。隐私属于重要的人权，未经许可任何单位或个人都不能擅自收集、处理、披露公民的个人数据与信息。建立在以人工智能技术与海量的政府和公民信息基础之上的数字政府使以往信息收集许可制度受到冲击。已有隐私信息收集许可，为使用者明示同意条件下的个人信息收集，人工智能技术打破了使用者所同意的个人信息收集的界限[26]。这是因为人工智能技术不但可以从公共数据中推导出个人信息，而且还可以从个人信息中推导出

个人的社会关系。同时，依托智能技术采集公民数据和生物信息时，如果缺乏相应的规范，可能造成了对公民隐私的侵害。

3　人工智能赋能数字政府建设风险的应对策略

人工智能赋能数字政府建设的过程中，弱化了政府建设对“人”的关系以及对行为的控制。因此，需要从人工智能在数字政府建设中的技术规范、科学审查、责任分配等方面着手，同时树立敏捷治理（Agile Governance）理念，以应对人工智能赋能数字政府建设的风险。四个应对策略与国家、政府、社会、个人层面安全风险的关系结构，如图 1 所示。

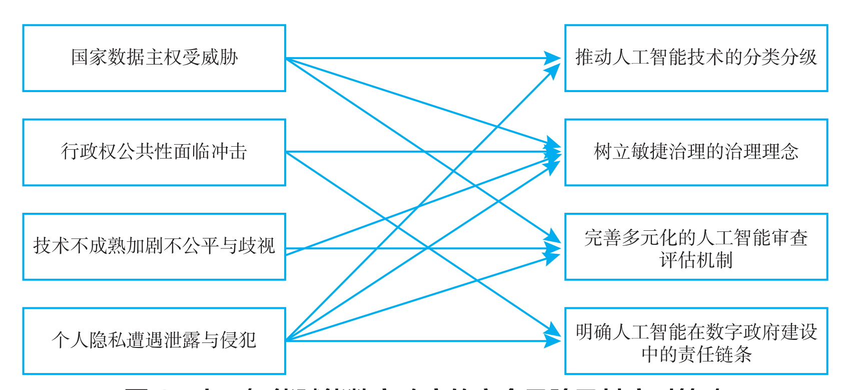


图 1　人工智能赋能数字政府的安全风险及其应对策略

3.1　技术规范：推动人工智能技术的分类分级

人工智能赋能数字政府建设，需要超越以技术为核心的效能提升观念，促进工具理性与价值理性的有效统一。推动人工智能技术的分类分级，是人工智能治理的源头节点，对人工智能在数字政府建设应用中的风险具有预防作用。作为一种技术规范，人工智能技术分类分级旨在根据人工智能的技术成熟度、应用场景、潜在风险等因素设定配套的审查、监管机制与责任承担方式。对于人工智能技术的分类分级，我国《生成式人工智能服务管理暂行办法》[27] 已经提出了分类分级监管的思路。考虑到人工智能技术的迭代问题，如果只对生成式人工智能进行分类分级无法适应人工智能技术的发展和应用。基于此，可以参照欧盟的做法，依据人工智能风险的大小将其分为不可接受风险、高风险、有限风险。不可接受风险类型的人工智能可能会侵犯基本人权、危及人类的生命安全，应当被禁止。高风险类型的人工智能流入市场前须经过严格的强制性审查、评估，并自觉接受全流程风险管理。有限风险类型的人工智能所受约束最小，只需履行一定的透明度义务，满足公民的知情权，使人们规避特定的被操纵风险[28]。在具体数字政府建设中，需要在分类分级的基础上进一步明确人工智能在数字政府建设中的使用范围与领域。具体来说，可以从人工智能在数字政府建设中行为的确定性高低、难易程度、裁量度高低、受益和负担行为划分、对于当事人权益影响等进行考量，进而将确定性程度高的行为，事实清楚、证据通过技术容易固化、法律规定明确的行为，羁束行政行为，给付行政行为，不影响当事人权益的行为确定为人工智能嵌入数字政府建设的范围。

3.2 理念赋能：树立敏捷治理的治理理念

敏捷治理经由软件开发领域到公共管理领域。敏捷治理理念的核心是以一种持续准备的状态，快速适应社会生活状况，主动或被动地接受变化并从中学习，为满足用户需求作出贡献，旨在为实现公共目标，接受利益相关方，运用快速、灵活、有效的方式适应不断变化的环境与回应公众需求[29]。人工智能赋能数字政府作为新兴领域、重点领域，其潜在风险大且具有高度不确定性，传统治理范式难以有效应对。敏捷治理具有区别于风险治理的全面性、适应性、灵活性等特征[30]，契合其潜在的治理需求，同时与我国在新兴技术领域采取的包容审慎监管不谋而合。首先，敏捷治理关注治理对象的全过程，人工智能赋能数字政府于辅助决策、数据收集、政民互动等各个阶段都可能产生未知风险，唯有对人工智能应用的全过程实施监测，才能及时发现风险并辅以治理措施。其次，敏捷治理对于场景变化具有极强的适应能力，当前对于人工智能应用数字政府的风险治理还没有完全成熟、可靠的经验，而敏捷治理无需借鉴经验，可以凭借其强大的适应性与学习能力对人工智能风险治理作出回应。最后，敏捷治理可以满足不同主体利益，数字行政对个人、企业、政府的利益均会有所涉及，如个人隐私保护、商业秘密保护、行政管理效率、公共服务优化等。个人、企业将基于不同的价值追求参与监督过程，这种多元化的监督体系有利于及时发现和处置各类风险。

3.3 科学审查：完善多元化的人工智能审查评估机制

面对科学技术的复杂性和不确定性，人工智能嵌入数字政府建设需要进行预防性审查。当前我国对于人工智能的评估规范仍需进一步完善。首先，赋予专责监管部门评估人工智能技术的能力。近年来，政府绩效评估在我国日益受到关注，并与政府责任制相结合，成为了考核政府绩效管理的重要标准。在该绩效管理体系下，各级政府依托相关职能部门或第三方专业机构，形成了既具地方特色又符合统一标准的绩效评估方式。对人工智能评估机制的建设可以这种成熟、系统的评估方式为基础，并根据人工智能应用的特性施以一定变化。其次，推动企业内部审查。企业是人工智能技术研发者和应用者，可以通过推进科技创新，建章立制，进行内部审查。推动技术创新有助于填补人工智能技术漏洞。以联邦学习为例，在科研、市场、法律等激励下，联邦学习可解决数据隔离和数据隐私保护问题，通过建立数据“联盟”，让所有参与方都能受益，推动技术整体不断进步[31]。再次，强化社会组织的监督。面对“政府多部门混治”局面，社会组织凭借其专业性、灵活性、纽带性优势[32]，成为参与人工智能风险治理的重要一环。一方面，利用社会组织灵活创新的优势，根据人工智能的审查需求，吸纳具备相应技术能力的数字人才，形成具有数字专业特色的组织队伍，通过长期参与人工智能的审查工作，不断提高其审查技能的专业性，更好地履行审查职责；另一方面，社会组织可以履行监督职责，通过对科技伦理、算法透明可解释、正当使用等风险防治环节进行监督，激活社会组织在风险预防阶段的主体作用。最后，完善公民数字权利体系。公众参与是现代社会一种常见的民主方式，是民主理论和行政控权理论发展的产物。尽管我国《中华人民共和国民法典》《中华人民共和国个人信息保护法》为公民维护数字权利提供了法律依据，但与数字科技发展相比，公民数字权利体系仍需进一步完善。未来，可进一步赋予公民广泛的数字知情权、参与权，以此为基础建立便利的参与机制。

3.4 责任分配：明确人工智能在数字政府建设中的责任链条

人工智能有着极强的技术性，可以依靠强大的算力揭示难以捕获的数据关联性、事件规律性，且其中的逻辑架构难以被公众所理解，但这并不意味着行政机关与人工智能技术服务提供者享有责任豁免权，反而应该完善现有的追责链条，使之与技术创新力度相匹配。从责任链条出发，政府需要承担的责任包括引进人工智能的决策责任和人工智能运行的监督责任。具体来讲，一是政府引入人工智能的安全审查义务。各地政府在引入人工智能技术时，必须强化提供优质公共服务的责任，审慎考察人工智能特定的风险，尽可能从源头切断风险，否则政府将可能承担审查不当的行政责任。二是政府的决策责任与监督责任。尽管人工智能技术下算法决策是自动化决策的一种形式，但如何使用、何时使用都是行政机关内部先行决策的结果。也就是说，人工智能技术下算法决策还是受到行政机关的控制。因而，行政机关必须重视算法本身存在的偏见问题，在使用过程中要履行监督责任，不然一旦因为决策失误或监管不力等原因造成了危害，行政机关难辞其咎。人工智能以算法为核心，算法根本上属于“产品”，若运行过程算法模型出现运算逻辑、应用数据等瑕疵，人工智能服务提供者应当承担产品违约责任，行政机关可以向服务提供方追责，具体包括赔偿损失、单方解除合同、纳入政府采购黑名单等。责任链条的末端，涉及的是人工智能侵权责任的认定问题。对此，我国的司法实践主要依照《中华人民共和国民法典》《中华人民共和国产品质量法》规定的产品责任以及使用人的过错责任，将人工智能的生产者、销售者或使用者认定为责任主体。尽管有些学者认为此种路径存在不足，但整体而言这一路径具有救济损害、预防损害和警示损害的调整功能，并非仅具有权宜性，未来在产品责任的适用上，需要调整、细化人工智能侵权责任的责任主体、归责原则、缺陷认定、赔偿方法以及举证责任等规则[33]。

4 结 语

人工智能赋能数字政府建设的过程是自动化行政广泛应用的过程。作为经验与算法加持下的产物，自动化决策在促进数字政府建设的同时，衍生出一系列潜在的风险。这些风险在国家层面表现为数据主权受到威胁，在政府层面体现在行政权公共性面对冲击，在社会层面呈现为不公平与歧视的风险加剧，在个人层面表现为公民隐私可能遭遇侵犯。面对人工智能赋能数字政府建设带来的以上风险，建议从推动人工智能技术的分类分级、树立敏捷治理的治理理念、完善多元化的人工智能审查评估机制、明确人工智能在数字政府建设中的责任链条四个层面予以应对。

【参考文献】

[1] 张翔. “复式转型”：地方政府大数据治理改革的逻辑分析［ J ］.中国行政管理，2018（12）：37–41.

[2] 张茜茜，涂群. 生成式人工智能袭来,数字政府市场格局正在发生巨变［ EB/OL ］.［ 2023–10–30 ］. https：//www.secrss.com/articles/54246.

[3] Calo R. Artificial intelligence policy: a primer and roadmap［ J ］. U.C.Davis Law Review, 2017, 51: 399–436.

[4] 王芳，张超，黄梅银，等. 数智赋能政府治理的理论与实践进展：一个跨学科学术会议综述［ J ］.图书与情报，2023（3）：126–135.

[5] 汪太贤，唐祎. 人工智能嵌入政府治理：算法图景、价值问题与回归路径［ J ］.中国科技论坛，2023

（2）：104-113.

［6］李良成，李雨青. 人工智能嵌入政府治理的风险及其规避［J］.华南理工大学学报（社会科学版），2021，23（5）：1-13.

［7］刘玮，王锋. 政务服务智能化创新的演化、风险与图景——基于场域视角的分析［J］. 电子政务，2024（2）：79-88.

［8］陈兵，董思琰. 生成式人工智能的算法风险及治理基点［J］. 学习与实践，2023（10）：22-31.

［9］陈礼，吕佩安. 数字政府治理中的类ChatGPT模型研究［J］. 征信，2023，41（10）：6-17.

［10］Autor D H. Why Are There Still So Many Jobs? The History and Future of Workplace Automation［J］.The Journal of Economic Perspectives, 2017, 19(3): 123-136.

［11］Turner J. Robot rules: regulating artificial intelligence［M］. PalgraveMacmillan,2019:253-280.

［12］Wright D, Raab C D. Constructing a surveillance impactassessment［J］.Computer Law and Security Review:The International Journal of Technology and Practice, 2012, 28(6): 613-626.

［13］Urs, Gasser, Virgilio, et al. A Layered Model for AI Governance［J］.IEEE Internet Computing, 2017, 21(6): 58-62.

［14］秦小建，周瑞文. 人工智能嵌入政府治理的探索及启示［J］. 国外社会科学，2022（2）：30-45.

［15］曾宇航，史军. 政府治理中的生成式人工智能：逻辑理路与风险规制［J］. 中国行政管理，2023，39（9）：90-95.

［16］张欣，高天书. 我国应用ChatGPT辅助行政决策算法危机治理的探索与思考［J］. 现代管理科学，2023（5）：81-88.

［17］刘绍宇. ChatGPT模型融入数字政府建设研究：风险、基础与路径［J］. 企业经济，2023，42（10）：95-104.

［18］范柏乃，林哲杨. 政府治理的“法治—效能”张力及其化解［J］. 中国社会学，2022（2）：162-184.

［19］阙天舒，王子. 数字经济时代的全球数据安全治理与中国策略［J］. 国际安全研究，2022（1）：130-154，158.

［20］蒋万胜，刘玲霞. 国际数字主权领域内的多元竞合与应对［J］. 中国特色社会主义研究，2023（5）：64-73.

［21］张夏恒. 类ChatGPT人工智能技术嵌入数字政府治理：价值、风险及其防控［J］. 电子政务，2023（4）：45-56.

［22］唐林垚. 遏制人工智能算法的公共妨害［N］. 法治时报，2020-01-07（07）.

［23］贾开. 人工智能与算法治理研究［J］. 中国行政管理，2019（1）：17-22.

［24］戚莹，高文英. 人工智能时代自动化行政的实践困境及规制路径［J］. 中国人民公安大学学报（社会科学版），2022，38（1）：67-75.

［25］余成峰. 信息隐私权的宪法时刻规范基础与体系重构［J］. 中外法学，2021，33（1）：32-56.

［26］王砼，康京涛，徐瑞泽. 法治政府数字化转型的逻辑、难点与对策［J］. 科学发展，2024（2）：65-72.

［27］中国网信网. 国家网信办等七部门联合公布《生成式人工智能服务管理暂行办法》［EB/OL］.［2024-06-05］. https：//www.cac.gov.cn/2023-07/13/c_1690898326795531.htm.

［28］European Commission. Proposal for a regulation of the european parliament and of the council laying down harmonised rules on ar-tificial intelligence (Artificial Intelligence Act) and amending certain union legislative acts［EB/OL］.［2024-06-05］.https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206.

［29］Mergel I, Gong Y, Bertot J. Agile government: systematic literature review and future research［J］.Government information quarterly, 2018, 35(2): 291-298.

［30］张凌寒，于琳. 从传统治理到敏捷治理：生成式人工智能的治理范式革新［J］.电子政务，2023

（9）：2–13.
［31］张吉豫. 构建多元共治的算法治理体系［J］. 法律科学（西北政法大学学报），2022，40（1）：115–123.
［32］张勤，钱洁. 促进社会组织参与公共危机治理的路径探析［J］. 中国行政管理，2010（6）：88–92.
［33］杨立新. 人工智能产品责任的功能及规则调整［J］. 数字法治，2023（4）：27–39.

Research on Security Risks and Response Strategies of AI-enabled Digital Government Construction

Kang Jingtao　Wang Tong

(School of Law, Xi'an University of Finance and Economics, Xi'an 710061, China)

Abstract:［**Purpose/Significance**］This paper carries out the risks of AI-enabled digital government construction and studies the risk response strategies in order to promote the high-quality development of China's digital government construction.［**Method/Process**］By using the methods of normative analysis and case study, this paper analyzes the risks and hidden dangers existing in the construction of AI-enabled digital government at the levels of country, government, society and individual.［**Result/Conclusion**］In response to the risks, this paper puts forward four levels of risk response strategies, which maily include promoting the classification and grading of AI technology, establishing the governance concept of agile governance, improving the diversified review and evaluation machanism for AI, and clarifying the responsibility chain of AI in the construction of digital government, with a view to providing new ideas for risk governance of AI-enabled digital government construction in China.

Keywords: Artificial intelligence(AI); Digital government construction; Data sovereignty; Social discrimination; Citizen privacy

（本文责编：孔青青）